EL054227599

EL365901262

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

* * * * *

GATED SEMICONDUCTOR ASSEMBLIES AND METHODS OF FORMING GATED SEMICONDUCTOR ASSEMBLIES

* * * * *

INVENTORS

Mark A. Helm
Mark Fischer
John T. Moore
Scott Jeffrey DeBoer

ATTORNEY'S DOCKET NO. MI22-845

EL979977228

GATED SEMICONDUCTOR ASSEMBLIES AND METHODS OF FORMING GATED SEMICONDUCTOR ASSEMBLIES

TECHNICAL FIELD

The invention pertains to gated semiconductor assemblies, such as, for example, erasable, programmable read-only memories (EPROMS), electrically erasable proms (EEPROMS), and flash EEPROMS.

BACKGROUND OF THE INVENTION

Read-only-memories (ROMs) are memories into which information is permanently stored during fabrication. Such memories are considered "non-volatile" as only read operations can be performed.

Each bit of information in a ROM is stored by the presence or absence of a data path from the word (access) line to a bit (sense) line. The data path is eliminated simply by insuring no circuit element joins a word and bit line. Thus, when the word line of a ROM is activated, the presence of a signal on the bit line will mean that a 1 is stored, whereas the absence of a signal indicates that a 0 is stored.

If only a small number of ROM circuits are needed for a specific application, custom mask fabrication might be too expensive or time consuming. In such cases, it would be faster and cheaper for users to program each ROM chip individually. ROMs with such capabilities are referred to as programmable read-only-memories (PROMs). In the first PROMs which were developed, information could only be programmed once into the construction and then could not be erased. In such

PROMs, a data path exists between every word and bit line at the completion of the chip manufacture. This corresponds to a stored 1 in every data position. Storage cells during fabrication were selectively altered to store a 0 following manufacture by electrically severing the word-to-bit connection paths. Since the write operation was destructive, once the 0 had been programmed into a bit location it could not be erased back to a 1. PROMs were initially implemented in bipolar technology, although MOS PROMs became available.

Later work with PROMs led to development of erasable PROMs. Erasable PROMs depend on the long-term retention of electric charge as the means for information storage. Such charge is stored on a MOS device referred to as a floating polysilicon gate. Such a construction differs slightly from a conventional MOS transistor gate. The conventional MOS transistor gate of a memory cell employs a continuous polysilicon word line connected among several MOS transistors which functions as the respective transistor gates. The floating polysilicon gate of an erasable PROM interposes a localized secondary polysilicon gate in between the continuous word line and silicon substrate into which the active areas of the MOS transistors are formed. The floating gate is localized in that the floating gates for respective MOS transistors are electrically isolated from the floating gates of other MOS transistors.

Various mechanisms have been implemented to transfer and remove charge from a floating gate. One type of erasable

programmable memory is the so-called electrically programmable ROM (EPROM). The charge-transfer mechanism occurs by the injection of electrons into the floating polysilicon gate of selected transistors. If a sufficiently high reverse-bias voltage is applied to the transistor drain being programmed, the drain-substrate "pn" junction will experience "avalanche" breakdown, causing hot electrons to be generated. Some of these will have enough energy to pass over the insulating oxide material surrounding each floating gate and thereby charge the floating gate. These EPROM devices are thus called *f*loating-gate, *a*valanche-injection MOS transistors (FAMOS). Once these electrons are transferred to the floating gate, they are trapped there. The potential-barrier at the oxide-silicon interface of the gate is greater than 3 eV, making the rate of spontaneous emission of the electrons from the oxide over the barrier negligibly small. Accordingly, the electronic charge stored on the floating gate can be retained for many years.

When the floating gate is charged with a sufficient number of electrons, channel function is inhibited. The presence of a 1 or 0 in each bit location is therefore determined by the presence or absence of a conducting floating channel gate in each program device.

Such a construction also enables means for removing the stored electrons from the floating gate, thereby making the PROM erasable. This is accomplished by flood exposure of the EPROM with strong ultraviolet light for approximately 20 minutes. The ultraviolet light

creates electron-hole pairs in the silicon dioxide, providing a discharge path for the charge (electrons) from the floating gates.

In some applications, it is desirable to erase the contents of a ROM electrically, rather than to use an ultraviolet light source. In other circumstances, it would be desirable to be able to change one bit at a time, without having to erase the entire integrated circuit. Such led to the development of electrically erasable PROMs (EEPROMs). Such technologies include MNOS transistors, floating-gate tunnel oxide MOS transistors (FLOTOX), textured high-polysilicon floating-gate MOS transistors, and flash EEPROMs. Such technologies can include a combination of floating gate transistor memory cells within an array of such cells, and a peripheral area to the array which comprises CMOS transistors.

A prior art EPROM device is described with reference to semiconductor wafer fragment 10 of Figs. 1-3. Fig. 1 is a top view of wafer fragment 10, and Figs. 2 and 3 are cross-sectional side views along the lines labelled X-X and Y-Y, respectively, in Fig. 1. Wafer fragment 10 comprises a substrate 12, having field oxide regions 14 formed thereover. Substrate 12 can comprise, for example, lightly doped monocrystalline silicon. To aid in interpretation of the claims that follow, the term "semiconductive substrate" is defined to mean any construction comprising semiconductive material, including, but not limited to, bulk semiconductive materials such as a semiconductive wafer (either alone or in assemblies comprising other materials thereon), and

semiconductive material layers (either alone or in assemblies comprising other materials). The term "substrate" refers to any supporting structure, including, but not limited to, the semiconductive substrates described above.

Field oxide regions 14 can comprise, for example, silicon dioxide. An active region 15 extends over and within substrate 12 between field oxide regions 14. A floating gate 16 and a control gate 18 are formed over the active region. Gates 16 and 18 can comprise, for example, conductively doped polysilicon.

Floating gate 16 is separated from substrate 12 by a tunnel oxide layer 20. Gates 16 and 18 are separated from one another by an insulative layer 22 which can comprise, for example, a combination of silicon dioxide and silicon nitride, such as the shown ONO construction wherein a silicon nitride layer 17 is sandwiched between a pair of silicon dioxide layers 19. The silicon nitride comprises Si_3N_4, although other forms of silicon nitride are known. Such other forms include silicon enriched silicon nitride layers (i.e., silicon nitride layers having a greater concentration of silicon than Si_3N_4, such as, for example, Si_4N_4). An advantage of silicon-enriched silicon nitride layers relative to Si_3N_4 is that the silicon-enriched silicon nitride layers frequently do not require separate, discrete antireflective coatings formed between them and a photoresist. However, silicon enriched silicon nitride is difficult to pattern due to a resistance of the material to etching. Silicon enriched silicon nitride layers are formed to have a substantially

homogenous composition throughout their thicknesses, although occasionally a small portion of a layer (1% or less of a thickness of the layer) is less enriched with silicon than the remainder of the layer due to inherent deposition problems.

Wafer fragment 10 further comprises silicon dioxide layers 24 and 26 extending along sidewalls of gates 16 and 18, and comprises a silicon dioxide layer 28 over control gate 18. Layers 24, 26 and 28 can electrically insulate gates 16 and 18 from other circuitry (not shown) that may be present on substrate 12.

The gate assembly shown in Figs. 1-3 can be formed as follows. Initially, a portion of substrate 12 within the active region is oxidized to form an oxide layer which will ultimately be patterned into tunnel oxide 20. Next, a polysilicon layer is formed over the silicon dioxide layer, with the polysilicon layer ultimately being patterned to form floating gate 16. An antireflective coating is formed over the polysilicon layer, and a layer of photoresist formed over the antireflective coating.

After the photoresist is formed, it is patterned by selectively exposing portions of the photoresist to light to render the portions either more soluble or less soluble in a solvent than portions which are not exposed to the light. The antireflective coating absorbs light that penetrates the photoresist to prevent such light from reflecting back to either constructively or destructively interfere with other light passing through the photoresist. The photoresist is then exposed to the solvent

to remove the more soluble portions of the photoresist and leave a patterned photoresist block over a portion of the polysilicon layer that is to become floating gate 16.

The patterned photoresist block protects the portion of the polysilicon layer it covers, while uncovered portions of the antireflective coating, polysilicon layer, and silicon oxide layers are removed with an etch. The portions of the polysilicon layer and oxide layer which remain are in the shape of floating gate 16 and tunnel oxide 20.

After the etch of the antireflective coating, polysilicon and oxide, the photoresist and antireflective coating are removed from over floating gate 16. The polysilicon of floating gate 16 is then exposed to oxygen under conditions which form a silicon dioxide layer over exposed surfaces of the polysilicon to create oxidized sidewalls 24 and 26, and a portion of insulative layer 22. Subsequently, layers of silicon nitride and silicon dioxide are provided to complete formation of insulative layer 22. Next, a second polysilicon layer is provided and patterned to form control gate 18. The second polysilicon layer is then exposed to oxygen to form silicon dioxide layers 24 and 26 at the sidewalls of control gate 18, and to form silicon dioxide layer 28 over a top of control gate 18.

Source and drain regions can be provided within active area 15 and operatively adjacent floating gate 16. The source and drain regions can be provided by implanting a conductivity enhancing dopant into

substrate 12 after forming floating gate 16 and before oxidizing sidewalls of floating gate 16.

A continuing goal in semiconductor device fabrication is to minimize the number of fabrication steps required to form a semiconductor device. Accordingly, it would be desired to eliminate one or more of the above-discussed steps in forming a gated semiconductor assembly.

SUMMARY OF THE INVENTION

In one aspect, the invention encompasses a method of forming a gated semiconductor assembly. A silicon nitride layer is formed over and against a floating gate. A control gate is formed over the silicon nitride layer.

In another aspect, the invention encompasses a method of forming a semiconductor assembly. A first material layer is formed over a substrate. A silicon nitride layer is formed over the first material layer. The silicon nitride layer comprises a first portion and a second portion elevationally displaced from the first portion. The first portion has a greater stoichiometric amount of silicon than the second portion. A photoresist layer is formed over the first material layer and the silicon nitride layer. The photoresist layer is patterned. The patterning comprises exposing portions of the layer of photoresist to light and utilizing the silicon nitride layer as an antireflective surface during the

exposing. The pattern is transferred from the patterned photoresist to the silicon nitride layer and the first material layer.

In yet another aspect, the invention encompasses a gated semiconductor assembly comprising a substrate, a floating gate over the substrate, a control gate over the floating gate, and an electron barrier layer between the floating gate and the control gate. The electron barrier layer comprises a silicon nitride layer. The silicon nitride layer comprises a first portion and a second portion elevationally displaced from the first portion. The first portion has a greater stoichiometric amount of silicon than the second portion.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of the invention are described below with reference to the following accompanying drawings.

Fig. 1 is a fragmentary, diagrammatic top view of a prior art gated semiconductor assembly

Fig. 2 is a diagrammatic, fragmentary, cross-sectional view of the Fig 1 gated semiconductor assembly along the line X-X of Fig. 1.

Fig. 3 is a diagrammatic, fragmentary, cross-sectional view of the Fig 1 gated semiconductor assembly along the line Y-Y of Fig. 1.

Fig. 4 is a diagrammatic, fragmentary, cross-sectional view of a semiconductor wafer fragment at a preliminary processing step of a method of the present invention, shown along an axis corresponding to line X-X of Fig. 1.

Fig. 5 is a view of the Fig. 4 wafer fragment shown along an axis corresponding to line Y-Y of Fig. 1.

Fig. 6 is a view of the Fig. 4 wafer fragment shown at a processing step subsequent to that of Fig. 4, and shown along an axis corresponding to line X-X of Fig. 1.

Fig. 7 is a view of the Fig. 6 wafer fragment shown along an axis corresponding to line Y-Y of Fig. 1.

Fig. 8 is a view of the Fig. 4 wafer fragment shown at a processing step subsequent to that of Fig. 6, and shown along an axis corresponding to line X-X of Fig. 1.

Fig. 9 is a view of the Fig. 8 wafer fragment shown along an axis corresponding to line Y-Y of Fig. 1.

Fig. 10 is a view of the Fig. 4 wafer fragment shown at a processing step subsequent to that of Fig. 8, and shown along an axis corresponding to line X-X of Fig. 1.

Fig. 11 is a view of the Fig. 10 wafer fragment shown along an axis corresponding to line Y-Y of Fig. 1.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

This disclosure of the invention is submitted in furtherance of the constitutional purposes of the U.S. Patent Laws "to promote the progress of science and useful arts" (Article 1, Section 8).

A method of forming a gated semiconductor assembly in accordance with the present invention is described with reference to

Figs. 4-11. Figs 4, 6, 8 and 10 are views of a semiconductor wafer fragment shown at sequential steps of a fabrication process, and shown along an axis corresponding to line X-X of Fig. 1. Figs. 5, 7, 9 and 11 are views of the processed wafer fragments of Figs. 4, 6, 8 and 10, respectively, along an axis corresponding to line Y-Y of Fig. 1.

Referring to Figs. 4 and 5, a semiconductor wafer fragment 40 comprises a substrate 42, and field oxide regions 44 formed over substrate 42. Substrate 42 and field oxide regions 44 can comprise the same compositions as substrate 12 and field oxide regions 14 discussed above in the "background" section. A portion of substrate 42 between field oxide regions 44 is defined as an active region.

A first silicon dioxide layer 50 is formed over the active region. Silicon dioxide layer 50 can be formed by, for example, exposing a silicon-comprising substrate 42 to oxygen at temperatures of, for example, at least 800° C.

A silicon-comprising floating gate layer 52 is formed over field oxide regions 44 and first silicon dioxide layer 50. Floating gate layer 52 can comprise, for example, amorphous silicon or polycrystalline silicon, and can be formed by, for example, chemical vapor deposition. The silicon of layer 52 is preferably doped with a conductivity-enhancing dopant to a concentration of greater than or equal to 1×10^{19} atoms/cm^3. The dopant can be, for example, provided *in situ* during the chemical vapor deposition process, or provided by implanting it into layer 52.

A silicon nitride layer 54 is formed over floating gate layer 52. In the shown preferred embodiment, silicon nitride layer 54 comprises a first portion 56 and a second portion 58, with one of portions 56 and 58 having a higher stoichiometric amount of silicon than the other of portions 56 and 58. Preferably, upper portion 58 will have a greater stoichiometric amount of silicon than will lower portion 56. An interface between portions 56 and 58 is illustrated with dashed line 57. In the shown embodiment, silicon nitride layer 54 is formed against floating gate layer 52. In other embodiments (not shown) an intervening silicon oxide layer can be formed between silicon nitride layer 54 and floating gate layer 52. Such intervening oxide layer can be formed by, for example, chemical vapor deposition or growth from the silicon of floating gate layer 52.

A preferred method of forming silicon nitride layer 54 is a chemical vapor deposition process. A silicon precursor gas and a nitrogen precursor gas are flowed into a reaction chamber at a first ratio to form portion 56 of silicon nitride layer 54, and then the ratio is changed to form portion 58. The silicon precursor gas can comprise, for example, SiH_2Cl_2 (dichlorosilane), and the nitrogen precursor gas can comprise, for example, NH_3 (ammonia). Example conditions for depositing silicon nitride from NH_3 and SiH_2Cl_2 comprise temperatures of from about 700° C to about 800° C, and pressures of from about 100 mTorr to about 1 Torr.

In a process wherein upper portion 58 is to have a greater stoichiometric amount of silicon than lower portion 56, the initial ratio of SiH_2Cl_2 to NH_3 flowed into a chemical vapor deposition can be, for example, about 0.33. Such ratio is flowed into the reaction chamber until first portion 56 is formed to a thickness of from about 50 Angstroms to about 500 Angstroms, and preferably to a thickness of about 75 Angstroms. The ratio of SiH_2Cl_2 to NH_3 of about 0.33 forms a first portion 56 having a stoichiometry of about Si_3N_4.

After forming first portion 56, the ratio of SiH_2Cl_2 to NH_3 is adjusted to be greater than 0.33 (such as, for example, about 6) to form upper portion 58. Upper portion 58 is preferably formed to a thickness of from about 50 Angstroms to about 500 Angstroms, preferably to a thickness of less than or equal to about 200 Angstroms, and more preferably to a thickness of less than or equal to about 100 Angstroms. Upper portion 58 preferably comprises a stoichiometry of Si_xN_y, wherein a ratio of x to y is at least 1. For example, upper portion 58 can comprise one or more of Si_4N_4, Si_7N_4 and $Si_{10}N_1$. If the ratio of SiH_2Cl_2 to NH_3 is about 6, upper portion 58 will have a stoichiometry of about Si_4N_4.

Preferably, portions 56 and 58 are formed in a common and uninterrupted deposition process. By "common deposition process" it is meant a deposition process wherein a wafer is not removed from a reaction chamber between the time that an initial portion of a silicon

nitride layer is formed and the time that a final portion of the silicon nitride layer is formed. By "uninterrupted deposition process" it is meant a process wherein the flow of at least one of the silicon precursor gas and the nitrogen precursor gas does not stop during the deposition process.

In a most preferred embodiment of the invention, floating gate layer 52 and silicon nitride layer 54 will be formed in a common and uninterrupted deposition process. Such uninterrupted deposition process can comprise, for example, flowing SiH_2Cl_2 into a chemical reaction chamber, without NH_3 being flowed into the chamber, to deposit a silicon-comprising floating gate layer 52 over substrate 42. Floating gate layer 52 is preferably formed to a thickness of from about 200 Angstroms to about 2000 Angstroms. After formation of floating gate layer 52, the SiH_2Cl_2 flow is maintained (although it may be reduced or increased) and a flow of NH_3 is initiated in the chamber to form first portion 56 of silicon nitride layer 54. The ratio of SiH_2Cl_2 to NH_3 flowing within the reaction chamber is then altered to form second portion 58 of silicon nitride layer 54.

After formation of silicon nitride layer 54, a patterned photoresist layer 60 is formed over silicon nitride layer 54. Patterned photoresist 60 is formed as follows. A photoresist material is provided over silicon nitride layer 54. The photoresist material is then exposed to a patterned beam of light to render portions of the material other than those of patterned layer 60 more soluble in a solvent than is the

material of patterned layer 60. The solvent is then utilized to remove the more soluble portions and leave patterned layer 60.

Silicon nitride layer 54 can be utilized as an antireflective layer during exposure of the photoresist material to light. Specifically, it is observed that a refractive index of a silicon nitride layer increases as a stoichiometric amount of silicon increases within the layer. For instance, it is observed that Si_4N_4 has a reactive index of 2.2, Si_7N_4 has a refractive index of 2.5, $Si_{10}N_1$ has a refractive index of 3.0, and Si_3N_4 has a refractive index of only 2.0. A material is typically considered a suitable antireflective coating material if it has a refractive index of at least 2.2. Accordingly, the portions of silicon nitride layer 54 having a stoichiometry of Si_xN_y, where an x is at least equal to y, can be suitable antireflective materials.

Referring to Figs. 6 and 7, a pattern from patterned photoresist layer 60 (Figs. 4 and 5) is transferred to layers 54 and 52 to pattern layers 54 and 56 into a floating gate stack 66. The pattern of photoresist layer 60 can be transferred to layers 52 and 54 by etching portions of layers 52 and 54 which are not covered by photoresist layer 60. A suitable etch can comprise, for example, a plasma-enhanced etch utilizing NF_3 and HBr.

Photoresist layer 60 (Figs. 4 and 5) is removed from over silicon nitride layer 54. Subsequently, a layer of silicon dioxide 64 is grown over gate stack 66. Silicon dioxide layer 64 is formed along a sidewall and over a top surface of gate stack 66. Silicon dioxide layer 64 can

be formed by, for example, growth from silicon of layers 52 and 54, or by chemical vapor deposition. Growth of silicon dioxide layer 64 can be accomplished by exposing gate stack 66 to an atmosphere comprising oxygen atoms at a temperature of at least about 500° C.

Referring to Figs. 8 and 9, a control gate layer 74 is formed over gate stack 66 and substrate 42, and a patterned photoresist mask 76 is formed over control gate layer 74. Control gate layer 74 can comprise, for example, conductively doped amorphous silicon or polycrystalline silicon, and can be formed by, for example, chemical vapor deposition.

Referring to Figs. 10 and 11, a pattern is transferred from mask 76 (Figs. 8 and 9) to control gate layer 74 to form layer 74 into a control gate over gate stack 66. The pattern can be transferred, with, for example, a plasma-enhanced etch utilizing NF_3 and HBr.

After formation of the control gate, an oxide layer 80 is formed over exposed surfaces of layers 52, 54 and 74. Oxide layer 80 can be formed by, for example, growth from the silicon of the control gate, or chemical vapor deposition.

Source and drain diffusion regions 72 are formed adjacent gate stack 66. Source and drain diffusion regions 72 can be formed by, for example, implanting a conductivity-enhancing dopant into substrate 42.

If one or both of floating gate layer 52 and control gate layer 74 comprise amorphous silicon, such layers are preferably converted to polycrystalline silicon in the gated semiconductor assembly of Figs. 10

and 11. Such conversion can occur by, for example, thermal processing of the layers at a temperature of at least about 700° C, and preferably from about 700° C to about 1100° C.

An advantage of the method of the present invention relative to prior art gated semiconductor assembly fabrication processes is that the method of the present invention can utilize an insulative material layer (54) as an antireflective surface during photolithographic processing of the insulative layer. Accordingly, the method of the present invention can eliminate a prior art utilization of a separate antireflective coating layer during patterning of an insulative layer over a floating gate construction. Another advantage of the method of the present invention is that it enables a common and uninterrupted deposition process to be utilized for formation of both a floating gate layer and an insulative layer over the floating gate layer.

As discussed above, it can be advantageous to have silicon nitride layer 54 comprise a portion having a stoichiometry of Si_xN_y, wherein x is greater than or equal to y, as such portion can be utilized as an antireflective layer. It is noted that it can also be advantageous to have silicon nitride layer 54 comprise a portion with a stoichiometry of Si_xN_y, wherein x is less than y, because such portion can be easier to etch than a portion having a greater stoichiometric amount of silicon. Accordingly, by having both types of portions between silicon nitride layer 54, the layer can be utilized as an antireflective material, and yet can be relatively easily removed when patterned.

In compliance with the statute, the invention has been described in language more or less specific as to structural and methodical features. It is to be understood, however, that the invention is not limited to the specific features shown and described, since the means herein disclosed comprise preferred forms of putting the invention into effect. The invention is, therefore, claimed in any of its forms or modifications within the proper scope of the appended claims appropriately interpreted in accordance with the doctrine of equivalents.